### ****Ejercicio Propuesto 8 (Resuelto)****

**Se realiza un estudio sobre el efecto que produce la descarga de aguas residuales de una planta sobre la ecología del agua natural de un río. En el estudio se utilizaron dos lugares de muestreo. Un lugar está aguas arriba del punto en el que la planta introduce aguas residuales en la corriente; el otro está aguas abajo. Se tomaron muestras durante un periodo de cuatro semanas y se obtuvieron los datos sobre el número de diatomeas halladas. Los datos se muestran en la tabla adjunta:**

|  |  |
| --- | --- |
|  | **Semanas** |
| **Lugar** | **Semana 1** | **Semana 2** | **Semana 3** | **Semana 4** |
| **Aguas****Arriba** | **78 94****43 58** | **620 760****420 913** | **204 333****98 89** | **890 655****763 562** |
| **Aguas****abajo** | **79 87****145 522** | **546 652****76 94** | **45 69****59 62** | **254 86****789 267** |

**Responder a las siguientes cuestiones:**

1. **Identificar el diseño adecuado a este experimento, escribir el modelo matemático y explicar los distintos elementos que intervienen.**
2. **Estudiar si la semana y el lugar son factores determinantes en el número de diatomeas halladas en el agua del río. ¿Hay posibilidad que una semana sea más recomendable en un lugar del río en concreto y no lo sea en el otro lugar?**
3. **Estudiar en qué semana se producen menos contaminación en el río, utilizando el método de Duncan.**
4. **Estudiar en qué lugar del río se producen menos diatomeas.**

**Solución:**

**Responder a las siguientes cuestiones:**

**1. Identificar el diseño adecuado a este experimento, escribir el modelo matemático y explicar los distintos elementos que intervienen.**

Es un diseño factorial dedos factores con replicación, la variable respuesta $y\_{ijk}$ viene descrita por la siguiente ecuación

Modelo estadístico del diseño factorial de dos factores con replicación

$$y\_{ijk }=μ+τ\_{i }+β\_{j}+(τβ)\_{ij}+u\_{ijk} , i=1,2 ; j=1,2,3,4; k=1,2,3,4$$

* ***yijk***: Número de diatomeas halladas
* ***µ*:** Efecto constante, común a todos los niveles de los factores, denominado media global.
* ***τi***: Efecto medio producido por lugar de muestreo *i*, (∑iτi = 0).
* ***βj***: Efecto medio producido por la semana *j*, (∑j βj  = 0).
* ***(τβ)ij***: Efecto medio producido por la interacción entre el lugar de muestreo i y la semana j,   (∑i (τβ)ij = ∑j (τβ)ij = 0).
* ***uijk***: Vv aa. independientes con distribución N(0,σ).
* **Variable respuesta: *Número de diatomeas halladas***
* **Factor:** ***Lugar de muestreo*** (dos niveles).
* **Factor:** ***Semana*** (cuatro niveles).
* Ambos factores de **efectos fijos.**
* **Tamaño del experimento:** Número total de observaciones (32).

Para realizar este supuesto en *R* debemos introducir primero los datos de forma correcta. Podemos introducir los datos directamente en *R* de forma manual o introducirlos previamente en un archivo de texto o Excel y leerlos en *R*.

En este caso lo hacemos en un archivo de texto:

* **Variable respuesta: n\_diatomeas**
* **Factor**: **semana**que tiene cuatro niveles. Es un factor de **efectos fijos** ya que viene decidido que niveles concretos se van a utilizar.
* **Factor**: ***lugar*** con dos niveles y de **efectos fijos**.
* **Tamaño del experimento:** Número total de observaciones (32).

n\_diatomeas semana lugar

78 Semana1 Aguas\_arriba

94 Semana1 Aguas\_arriba

620 Semana2 Aguas\_arriba

760 Semana2 Aguas\_arriba

204 Semana3 Aguas\_arriba

333 Semana3 Aguas\_arriba

890 Semana4 Aguas\_arriba

655 Semana4 Aguas\_arriba

43 Semana1 Aguas\_arriba

58 Semana1 Aguas\_arriba

420 Semana2 Aguas\_arriba

913 Semana2 Aguas\_arriba

98 Semana3 Aguas\_arriba

89 Semana3 Aguas\_arriba

763 Semana4 Aguas\_arriba

562 Semana4 Aguas\_arriba

79 Semana1 Aguas\_abajo

87 Semana1 Aguas\_abajo

546 Semana2 Aguas\_abajo

652 Semana2 Aguas\_abajo

45 Semana3 Aguas\_abajo

69 Semana3 Aguas\_abajo

254 Semana4 Aguas\_abajo

86 Semana4 Aguas\_abajo

145 Semana1 Aguas\_abajo

522 Semana1 Aguas\_abajo

76 Semana2 Aguas\_abajo

94 Semana2 Aguas\_abajo

59 Semana3 Aguas\_abajo

62 Semana3 Aguas\_abajo

789 Semana4 Aguas\_abajo

267 Semana4 Aguas\_abajo

> setwd("C:/Users/Usuario/Desktop/Datos") \*directorio de trabajo donde estén los datos

> propuesto8<-read.table("propuesto8.txt", header = TRUE)

> propuesto8

 n\_diatomeas semana lugar

1 78 Semana1 Aguas\_arriba

2 94 Semana1 Aguas\_arriba

3 620 Semana2 Aguas\_arriba

4 760 Semana2 Aguas\_arriba

5 204 Semana3 Aguas\_arriba

6 333 Semana3 Aguas\_arriba

7 890 Semana4 Aguas\_arriba

8 655 Semana4 Aguas\_arriba

9 43 Semana1 Aguas\_arriba

10 58 Semana1 Aguas\_arriba

11 420 Semana2 Aguas\_arriba

12 913 Semana2 Aguas\_arriba

13 98 Semana3 Aguas\_arriba

14 89 Semana3 Aguas\_arriba

15 763 Semana4 Aguas\_arriba

16 562 Semana4 Aguas\_arriba

17 79 Semana1 Aguas\_abajo

18 87 Semana1 Aguas\_abajo

19 546 Semana2 Aguas\_abajo

20 652 Semana2 Aguas\_abajo

21 45 Semana3 Aguas\_abajo

22 69 Semana3 Aguas\_abajo

23 254 Semana4 Aguas\_abajo

24 86 Semana4 Aguas\_abajo

25 145 Semana1 Aguas\_abajo

26 522 Semana1 Aguas\_abajo

27 76 Semana2 Aguas\_abajo

28 94 Semana2 Aguas\_abajo

29 59 Semana3 Aguas\_abajo

30 62 Semana3 Aguas\_abajo

31 789 Semana4 Aguas\_abajo

32 267 Semana4 Aguas\_abajo

A continuación debemos transformar tanto la columna de los tratamiento como la de los bloques en un factor para podemos realizar los cálculos posteriores adecuadamente.

> propuesto8$semana <- factor(propuesto8$semana)

> propuesto8$semana

 [1] Semana1 Semana1 Semana2 Semana2 Semana3 Semana3 Semana4 Semana4 Semana1

[10] Semana1 Semana2 Semana2 Semana3 Semana3 Semana4 Semana4 Semana1 Semana1

[19] Semana2 Semana2 Semana3 Semana3 Semana4 Semana4 Semana1 Semana1 Semana2

[28] Semana2 Semana3 Semana3 Semana4 Semana4

Levels: Semana1 Semana2 Semana3 Semana4

> propuesto8$lugar <- factor(propuesto8$lugar)

> propuesto8$lugar

 [1] Aguas\_arriba Aguas\_arriba Aguas\_arriba Aguas\_arriba Aguas\_arriba

 [6] Aguas\_arriba Aguas\_arriba Aguas\_arriba Aguas\_arriba Aguas\_arriba

[11] Aguas\_arriba Aguas\_arriba Aguas\_arriba Aguas\_arriba Aguas\_arriba

[16] Aguas\_arriba Aguas\_abajo Aguas\_abajo Aguas\_abajo Aguas\_abajo

[21] Aguas\_abajo Aguas\_abajo Aguas\_abajo Aguas\_abajo Aguas\_abajo

[26] Aguas\_abajo Aguas\_abajo Aguas\_abajo Aguas\_abajo Aguas\_abajo

[31] Aguas\_abajo Aguas\_abajo

Levels: Aguas\_abajo Aguas\_arriba

2. **Estudiar si la semana y el lugar son factores determinantes en el número de diatomeas halladas en el agua del río. ¿Hay posibilidad que una semana sea más recomendable en un lugar del río en concreto y no lo sea en el otro lugar?**

Primero vamos a obtener la tabla ANOVA

> mod1 <- aov(n\_diatomeas~ semana + lugar + semana \* lugar, data = propuesto8 )

> mod1

Call:

 aov(formula = n\_diatomeas ~ semana + lugar + semana \* lugar,

 data = propuesto8)

Terms:

 semana lugar semana:lugar Residuals

Sum of Squares 1236724.8 235984.5 330818.2 914922.0

Deg. of Freedom 3 1 3 24

Residual standard error: 195.2479

Estimated effects may be unbalanced

> summary(mod1)

 Df Sum Sq Mean Sq F value Pr(>F)

semana 3 1236725 412242 10.814 0.00011 \*\*\*

lugar 1 235984 235984 6.190 0.02018 \*

semana:lugar 3 330818 110273 2.893 0.05616 .

Residuals 24 914922 38122

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Son significativos los efectos de los dos factores, pero no es significativo la interacción. Ambos son factores determinantes en el número de diatomeas halladas en el agua del río.

Para determinar si hay algún día de la semana más recomendable que otro, vamos a aplicar un contraste de comparaciones múltiples

Recordar que hay que cargar e instalar el paquete agricolae

> library(agricolae)

> LSD.test(mod1,"semana", p.adj="bonferroni", console=TRUE)

Study: mod1 ~ "semana"

LSD t Test for n\_diatomeas

P value adjustment method: bonferroni

Mean Square Error: 38121.75

semana, means and individual ( 95 %) CI

 n\_diatomeas std r LCL UCL Min Max

Semana1 138.250 157.90933 8 -4.222082 280.7221 43 522

Semana2 510.125 299.51887 8 367.652918 652.5971 76 913

Semana3 119.875 99.41319 8 -22.597082 262.3471 45 333

Semana4 533.250 295.20344 8 390.777918 675.7221 86 890

Alpha: 0.05 ; DF Error: 24

Critical Value of t: 2.875094

Minimum Significant Difference: 280.6781

Treatments with the same letter are not significantly different.

 n\_diatomeas groups

Semana4 533.250 a

Semana2 510.125 a

Semana1 138.250 b

Semana3 119.875 b

El número de diatomeas (533.250) es mayor en la semana 4 y menor en la semana 3 (119.875). Hay dos subgrupos diferenciados. Un subgrupo está formado por las semanas 2 y 4 y el otro subgrupo por las semanas 1 y 3.

> SNK.test(mod1,"lugar", console=TRUE)

Study: mod1 ~ "lugar"

Student Newman Keuls Test

for n\_diatomeas

Mean Square Error: 38121.75

lugar, means

 n\_diatomeas std r Min Max

Aguas\_abajo 239.50 245.9295 16 45 789

Aguas\_arriba 411.25 324.0622 16 43 913

Alpha: 0.05 ; DF Error: 24

Critical Range

 2

142.4721

Means with the same letter are not significantly different.

 n\_diatomeas groups

Aguas\_arriba 411.25 a

Aguas\_abajo 239.50 b

Se produce menos contaminación en aguas abajo, con un número de diatomeas de 239.50. Y como hemos dicho anteriormente hay diferencias significativas entre el número de diatomeas que se producen en aguas arribas del rio y las que se producen en aguas abajo del rio

**3. Estudiar en qué semana se producen menos contaminación en el río, utilizando el método de Duncan.**

> (duncan=duncan.test(mod1, "semana" , group = T))

$statistics

 MSerror Df Mean CV

 38121.75 24 325.375 60.00704

$parameters

 test name.t ntr alpha

 Duncan semana 4 0.05

$duncan

 Table CriticalRange

2 2.918793 201.4859

3 3.065610 211.6208

4 3.159874 218.1279

$means

 n\_diatomeas std r Min Max Q25 Q50 Q75

Semana1 138.250 157.90933 8 43 522 73.00 83.0 106.75

Semana2 510.125 299.51887 8 76 913 338.50 583.0 679.00

Semana3 119.875 99.41319 8 45 333 61.25 79.0 124.50

Semana4 533.250 295.20344 8 86 890 263.75 608.5 769.50

$comparison

NULL

$groups

 n\_diatomeas groups

Semana4 533.250 a

Semana2 510.125 a

Semana1 138.250 b

Semana3 119.875 b

attr(,"class")

[1] "group"

El número de diatomeas (533.250) es mayor en la semana 4 y menor en la semana 3 (119.875)

Editor de R

setwd("C:/Users/Usuario/Desktop/Datos")

propuesto8<-read.table("propuesto8.txt", header = TRUE)

propuesto8

propuesto8$semana <- factor(propuesto8$semana)

propuesto8$semana

propuesto8$lugar <- factor(propuesto8$lugar)

propuesto8$lugar

mod1 <- aov(n\_diatomeas~ semana + lugar + semana \* lugar, data = propuesto8 )

mod1

summary(mod1)

library(agricolae)

LSD.test(mod1,"semana", p.adj="bonferroni", console=TRUE)

SNK.test(mod1,"lugar", console=TRUE)

(duncan=duncan.test(mod1, "semana" , group = T))